2017

Towards Massive Data and Sparse Data in Adaptive Micro Open Educational Resource Recommendation: A Study on Semantic Knowledge Base Construction and Cold Start Problem

Geng Sun
University of Wollongong, gs147@uowmail.edu.au

Tingru Cui
University of Wollongong, tingru@uow.edu.au

Ghassan Beydoun
University of Wollongong, beydoun@uow.edu.au

Shiping Chen
CSIRO ICT Centre, shiping.chen@csiro.au

Fang Dong
Southeast University, fdong@seu.edu.cn

See next page for additional authors

Abstract

Micro Learning through open educational resources (OERs) is becoming increasingly popular. However, adaptive micro learning support remains inadequate by current OER platforms. To address this, our smart system, Micro Learning as a Service (MLaaS), aims to deliver personalized OER with micro learning to satisfy their real-time needs.

Keywords

educational, open, micro, adaptive, problem, sparse, data, massive, start, cold, construction, base, knowledge, semantic, study, recommendation:, resource, towards

Disciplines

Engineering | Science and Technology Studies

Publication Details

Sun, G., Cui, T., Beydoun, G., Chen, S., Dong, F., Xu, D. & Shen, J. (2017). Towards Massive Data and Sparse Data in Adaptive Micro Open Educational Resource Recommendation: A Study on Semantic Knowledge Base Construction and Cold Start Problem. Sustainability, 9 (6), 898-1-898-21.

Authors

Geng Sun, Tingru Cui, Ghassan Beydoun, Shiping Chen, Fang Dong, Dongming Xu, and Jun Shen

Article

Towards Massive Data and Sparse Data in Adaptive Micro Open Educational Resource Recommendation: A Study on Semantic Knowledge Base Construction and Cold Start Problem

Geng Sun [1], Tingru Cui [1], Ghassan Beydoun[2], Shiping Chen [3], Fang Dong [4], Dongming Xu [5] and Jun Shen [1]

[1] School of Computing and Information Technology, University of Wollongong; gs147@uowmail.edu.au, tingru@uow.edu.au, jshen@uow.edu.au
[2] School of Systems, Management and Leadership, University of Technology Sydney; Ghassan.Beydoun@uts.edu.au.
[3] Data61, CSIRO, Shiping.Chen@data61.csiro.au
[4] School of Computer Science and Engineering, Southeast University, fdong@seu.edu.cn
[5] UQ Business School, The University of Queensland, D.Xu@business.uq.edu.au
* Correspondence: gs147@uowmail.edu.au

Academic Editor: name
Received: date; Accepted: date; Published: date

Abstract: Micro Learning through open educational resources (OERs) is becoming increasingly popular. However, adaptive micro learning support remains inadequate by current OER platforms. To address this, our smart system, Micro Learning as a Service (MLaaS), aims to deliver personalized OER with micro learning to satisfy their real-time needs. In this paper, we focus on constructing a knowledge base to support the decision-making process of MLaaS. MLaas is built in a top-down approach. A conceptual graph based ontology construction is first developed. An educational data mining and learning analytic strategy is then proposed for the data level. The learning resource adaptation still requires learners' historical information. To compensate for the absence of this information initially (aka 'cold start'), we set up a predictive ontology-based mechanism. As the first resource is delivered to the beginning of a learner's learning journey, the micro OER recommendation is also optimized using a tailored heuristic.

Keywords: Adaptive Learning, Micro Open Learning, Educational Data Mining and Learning Analytics, Cold Start Problem

---

1. Introduction

In the information age, the development and dissemination of learning resources are booming at a much higher speed and wider range than their traditional shapes. People have shown increasing interests in getting access to online learning resources and getting involved in online learning activities especially via mobile devices [1]. Many leading universities opened up access to their courses. Indeed, access to open education resources (OERs) is exponentially increasing. This boom of OERs gains wider popularity in the entire higher and adult education sector, and also has attracted many researchers' attention, from educational, social, and computational views [2]. According to the latest statistics, millions of people have attended the virtual classroom of online open learning to access OERs, which are produced and updated on a daily basis. This leads to an emerging concomitant trend – that of open learning [3].

Open learning is recognized as a novel and effective learning method that can lead to a revolution of the traditional learning, distance learning and electronic learning (e-learning) which were widely used in the first decade of 21th century. Nevertheless, the current OER delivery still faces challenges and its sustained success remains in doubt. Recent studies actually suggest that massive open online courses (MOOCs), a common open learning environment, are currently suffering from low completion rate [4]. Most learners who enroll in MOOC courses end up dropping out. Educational professionals have focused much of their efforts on exploring open learning, OER and MOOC format as a regular, but the concomitant pedagogical innovations for mobile learning are yet to receive sufficient attention [5]. Indeed, there are many opportunities to improve open learning and OERs delivery.

Our previous studies [6, 7] demonstrated that micro learning is becoming a mainstream online learning mode, but coupled with mobile platforms the knowledge attainment process is often fragmented. Using mobile devices, learners are easily affected by their mood or environmental distractions [8]. In this paper, we will present our research on profiling micro learning processes through OERs and building a knowledge base to support the decision-making process of micro OER adaptation. This knowledge base will be built using a top-down approach. A construction of augmented ontologies oriented to micro open learning will be illustrated first, followed by a data processing strategy. Due to the brevity of learner information at the commencement of the micro open learning, we will introduce an ontological approach to technically address the cold start problem.

2. Background

2.1. Nature of OERs and Open Learning Delivery in Mobile Environment

Open learning is different from on-campus, e/m-learning mode. OERs are "digital learning resources offered online freely and openly to teachers, educators, students, and independent learners in order to be used, shared, combined, adapted, and expanded in teaching, learning and research" [3]. Open learning is the combination of informal learning and formal learning. Learners enjoy high flexibilities in online open learning because there is no strict time constraint for joining and quitting. Learners engaged in open learning are from different age groups and culture backgrounds with a wide range of geographic distribution.

Generally, OERs can be differentiated from MOOC and open courseware (OCW). Contrary to MOOC, OCW only offers course materials rather than entire course. In other words, OER can be structured (MOOC content) or unstructured (i.e. OCW), even both of them. OER providers and instructors have tried to promote their courses and affiliated educational products at full stretch. They have leveraged mobile learning (m-learning) for learners to easily participate in learning activities regardless of restrictions in time and location.

From another aspect, mobile learning activities in open learning normally consist of two sections: online learning and offline learning [9]. Since mobile learners can freely download materials into their mobile devices to view offline, they do not often stay on open learning platforms and attend virtual classrooms [10]. In fact, accessing OERs online is only a part of any learning; more tasks associated with any learning would require activities offline [11], such as data collection, data analysis, and report writing for an assignment. Logically, mobile open learning is through online systems that include guided and instructional materials, the transaction details and deliverable resources [12]. Hence, while learners are able to accomplish many open learning tasks offline, for some necessary procedures, such as data entry and work submission, they need to go back online to conduct these specific tasks.

2.2. Micro Learning

Micro learning refers to short-term learning activities on small learning units [13]. Its learning process can cover a time span from a few seconds (e.g. in mobile learning) to up to 15 minutes [14]. With mobile devices, learners normally accomplish learning mission in a short time period. According to prior study [14], micro learning can be defined by the assumption that short time span is needed to complete a relevant learning task. Hence, micro learning booms with the wide use of mobile devices, and it becomes a major learning means in mobile environment. Micro learning shares some similar characteristics with mobile learning as both of them are individually referable, self-contained, reusable and re-mixable [15].

Micro learning resources are available on-demand to facilitate just-in-time learning [16]. These small learning bytes cannot be learnt on-the-go, but also require less effort. They can aid quick assimilation, thus reducing the dependency on a fixed time slot or the need to take a large chunk of time out of learners' working day [13]. As micro learning evolves, micro-content delivery with a sequence of micro interactions enables users to learn without information overload [16]. Compared to traditional learning modes, now learners' overall efforts to go through an entire concept will proceed in a continuous, or even intermittent, way rather than a consecutive way [16].

3. Research Challenges and Design

3.1. Research Design

3.1.1 System Framework and Previous Work

In our previous studies [6, 7], we have discussed the popularity of adopting micro learning in accessing OER, especially through mobile devices. The necessity of improving existing mechanism of micro learning support has also been stated out. Having studied the present status of research and development of open learning and OERs, we are motivated to carry out a research to provide learners adaptive OERs by the means of micro learning in regards to their individualities. In other words, we are dedicated to tailor OERs into chunks with relatively short time length and allocate it to learners at the right timing. This approach was realized by a Software as a Service, Micro learning as a Service (MLaaS). Making use of it, optimally learners can easily complete the learning process by using their fragmented pieces of time. For example, a learner may spend 15 minutes to use mobile devices to learn a piece of MOOC course on his or her way home from work by train. In this case, an ideal course module delivered to him or her should be limited to the time length (e.g. 15 minutes) to ensure a micro but complete learning experience.

The framework of MLaaS is shown in Figure 1. As a data-rich system, MLaaS will be able to exploit detailed learner activity data for not only recommending what the next micro learning activity for a particular student should be, but also predicting how that student will perform with future learning content.

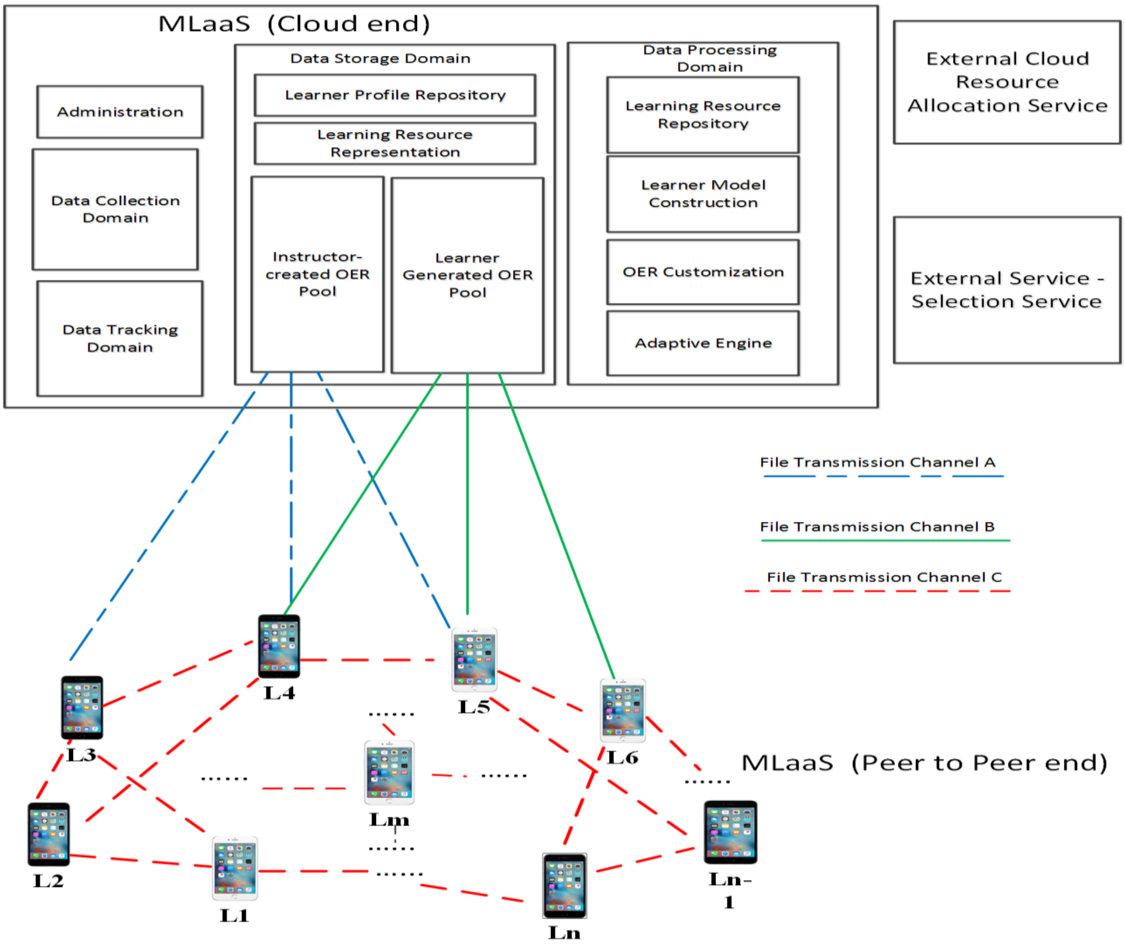

Figure 1. Topology of the Architecture of MLaaS

In our pilot work [17], we proposed a peer to cloud and peer-to-peer model for resource sharing and storage in service-oriented context. Such model can have higher upload and download speed than traditional cloud, user model or peer-to-sever-peer model, and be more robust to the failures of peers or servers in the cloud environment [17, 39]. Hence, we adopt this design and apply its concept as the topology of the new system for micro open learning.

The P2P sub-network of the proposed system is to conform with the nature of open learning, where various P2P learning occurs frequently and randomly. This p2p tier guarantees that P2P learning can be organized instantly, and the first hand resources can be shared and exchanged, regardless of the access of cloud.

From the top-down view, MLaaS borrow the cloud service to maximize the capability of hosting. Cloud part of the system consists of four domains: data tracking, data collection, data processing and data storage.

Functions of modules in MLaaS's cloud-end can be referred to prior works [5][18]. A noticeable future of the system is that there are three file transmission channels:

- A channel between learners and instructor-created OER pool in the cloud storage part (i.e. Channel A in Figure 1).
- A channel between learners and learner-generated OER pool in the cloud storage part (i.e. Channel B in Figure 1).
- A channel among all learners engaged in open learning (i.e. Channel C in Figure 1).

Once a learner shows his or her demand to carry out micro learning and sends such request from mobile device, OERs will be transmitted through one of the three channels.

Where the OERs actually come from the cloud resources pools (from which exact cloud nodes the OERs are retrieved and invoked) will be defined and externally supported by third-party service-selection and resource-allocation services from mainstream service providers. This problem has been well studied; typical solutions can be referred to the work reported in [19].

We reported the architecture and technical details of MLaaS in prior studies [18][20]. A comprehensive description is beyond the scope of this paper. It is worth noting that, MLaaS only produce micro OERs rather than OERs. That is to say, the normal OERs available online are collected by MLaaS and clustered in the OER pools, as shown in Figure 1. For this reason, despite the data collection mechanism that MLaaS owns, it shares some demographic and educational data with the platforms or providers where the OERs originally from. This helps in learner profiling, which will be introduced in the Section 5.1, even if a new learner registration in MLaaS is informal without sufficient demographic and educational data provided.

Therefore, the system framework is briefly introduced here as a background and we will skip to the focus of this paper, the ontology construction, data processing strategy and cold-start problem.

3.1.2 Research Problem Identification and Design

Given all decisions of micro OER adaptation are made by the Adaptive Engine, it acts as the core of the system. It consumes the results from all other services and transmits its output to the user interfaces straightforward. For this reason, the MLaaS is conceived to meet the standard of a data-rich system, and a knowledge base serves as its think tank. Basically, the knowledge base is constructed by a top-down approach, by making use of the semantics means, from the pattern level to data level. In other words, several ontologies are drawn at first, followed by a 'data processing' work. We attempt to combine a pattern and rule discovery process of micro learning, together with a survey of education literature for the features that could affect learning experience and outcomes in mobile environment [20]. In detail, the 'data processing' work involves all operations on data, from the very beginning to the end, such as entities extractions, relationship extractions, resolution disambiguation and so on.

As far as we have the overall system framework in place, consequently, we adopt the conceptual graph based approach to deal with the ontology construction [21, 40]. These graphs profile the features playing significant roles on an ongoing micro learning process and also depict how features were mutually affected and interrelated to each other. According to our design, the profiling procedure is carried out from two sides, the learner side and OER side.

While the profiling is proceeding forward, some new problems appear which go against the original intention while we were designing it. One of the most important problems is that the system, MLaaS, knows little about the learners, because either 'OER' or 'learner' is new to this emerging educational setting. This brings serious difficulties to launch the 'data processing' work. the profile construction is impossible with insufficient information about the learner at the commencement of open learning. Therefore, the learner profile cannot be fully identified with valid data.

In this paper, the research focuses on the knowledge base for the micro OER recommendation and delivery. Naturally, depending on the volume of retrievable data, this problem will be studied from two sides.

1) If a learner is well-known by the MLaaS, an educational data mining and learning analytics (EDM/LA) approach will be applied to his or her historical data to understand his or her learning patterns and preferences thereby a recommendation will be made well-grounded with respect to his or her personalized settings and particular surroundings.
2) If a learner is known scarcely by the MLaaS, (i.e., this is a new learner to the OER environment), this will be treated as a cold-start problem and tackled by filling in the gaps with predicted data, so that a recommendation will be made based on demographic information.
3) Information freshly generated along with the cold-start recommendation will be filled as the first version in a learners' profile.

3.2 Research Challenges

3.1.1.EDM/LA for Micro Learning

Student learning data collected by open learning systems are explored to develop predictive models by applying educational data mining methods that classify data or identify relationships. These models play a key role in building adaptive learning systems, in which adaptations or interventions based on the model's predictions can be used to change what students experience next, or even to recommend academic services to support their learning.

Analyzing these newly logged events requires new techniques to work with unstructured text and image data, data from multiple sources, and vast amounts of data ("big data"). Big data does not have a fixed size; any number assigned to define it would change as computing technology advances to handle more data. For example, Manyika et al. defines big data as "Datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze."[22]

At this cutting edge, EDM and LA are widely used in research. They are used to build models in several areas that can influence online learning systems. As its name implies, EDM is a state of art that applies the data mining techniques to educational data. It is concerned with many developing methods, and acts on exploring the unique types of data in educational settings. Using these methods, students and educational settings can be better understood [23]. To enable smart and adaptive micro learning for MOOC, EDM and LA are key concepts that we employ to build the basis of the dynamic learner model construction.

Generally, OCW data is locked away in independent data silos hosted by different OCW/OER providers. This makes it much less useful than it could be. It is difficult to develop tools for consume data from multiple silos. Searching OCW/OER across multiple silos means invoking each one's user interface, and receiving the results in separate groups. The presence of data silos makes accessing data and interoperability between repositories harder in several ways.

Browsing OERs also has the problem because each silo has its own organizational structure. Some silos have no way to link to a particular item, and so hinder the free flow of information. The presence of OCW silos impedes the interoperability, discovery, synthesis, and flow of knowledge. As a result, it is difficulty for teachers, students and self-learners to look resources – or sometimes they make decisions based on incomplete information. Linked data have the potential of create bridges between OCW data silos.

3.1.2. Cold Start Problem in Micro OER Delivery

In computer science literature, widely used adaptive recommendation methods generally consist of two main categories, i.e., memory-based and model-based algorithms [24]. Although they have been found in many successful cases of recommender systems, for example, Amazon online store, it is usually difficult to provide reliable recommendations due to the insufficiency of initial data of ratings or preferences. This leads to the occurrence of cold start problem. Commonly the cold start problem is triggered by three factors: new community, new item and new users.

The cold start problem becomes more severe in the open learning, especially in micro learning through OERs [14]. Both open learning and OER are relatively new products, which are emerging in the very recent years. Meanwhile, the followers of this novel trend, no matter new education pursuers or regular learners migrated from other online learning modes, are forming as a completely new community. On the other hand, the learning demands and expectations of learners engaged in open learning are much more practical than conventional university students. In other words, they are mostly self-regulated so that it is totally flexible for them to decide when to join or quit the online course at their own willingness, and switch among courses frequently [25]. Consequently, for OER providers, it is difficult to establish a model and update it accordingly for any individual learner because they do not have historical data in hand.

In micro open learning, or micro learning over OERs, it is very normal to find that learners take part in and deviate from the learning scenarios frequently, as well as turning on and off the learning activities at their own willingness. That is to say, the overall situations of micro learning vary all the way, from individual to individual. Moreover, it is very common that freshmen join into open learning or existing learners unfold a brand-new course learning profile, at any time. All in all, there are a large number of new learners in open learning; and a new learner usually initiates access of new learning resources; and learners who went through learning resources in a same branch of a discipline will form as a new community.

If treated carelessly, the cold start problem may lead to the loss of learners who are previously engaged in open learning and then decide to stop using the OER delivery system or adopting the learning mode [26]. The reasons behind the situation are mainly due to the lack of accuracy in the recommendations received in that first stage, in which the learners have not yet cast a significant number of votes or rating to feed the recommender systems. Basically, the sparsity of data affects the user satisfaction and then it can further affect the user acceptance of the new open learning mode.

3.3 Contribution

Extended from our published works [18][27], which presented the general framework and educational background of the micro open learning, this research will provide innovative deliverables by the following means:

- A top-down processing of semantic knowledge base building
- Conceptual graph based ontology constructions for the pattern level
- A data sources documentation and data processing strategy for the data level.
- A complete ontology-based mechanism to tackle cold start problem.

4. Conceptual graph Based Ontology Construction for Micro Open Learning and Proposed Data Processing Strategy

4.1 Conceptual Graph based Ontology Construction

Naturally a workable knowledge base has a two-tier structure, a pattern level at the top and data level at the bottom [28]. For the pattern level, the ontologies are constructed based on conceptual graphs, as we briefly introduced in the Section 3.1.1. By this means, the ontologies represent a formal way of the data processing workflow, and can drive the data processing with a priori knowledge and reduce the search space [29, 40].

By accomplishing a comprehensive survey on literature in the fields of pedagogy, psychology, e-learning and mobile learning, we sorted out features that might play key roles in the micro open learning experience and achievement. These conceptual graphs also represent how features affected and interrelated with each other in the ongoing micro open learning process. It will be introduced in the subsequent sections, 4.1.1 and 4.1.2, from the OER (item) side and learner (user) side, respectively.

4.1.1 Augmented Micro OER Ontology

From the item-based view, we deepen the sights into the micro learning environment in particular and, for this reason, the general ontology of OER is augmented to adapt the needs of micro learning.

In the augmented micro OER ontology, an annotation of a micro OER is self-describing with metadata exploring its educational parameters, such as typology (video, audio, text, etc.), type of interaction (expositive, active, mixed, two-way), didactic model (e.g. inductive, deductive, learning by doing, etc.), and non-functional attributes, such as QoS, semantic density and so on [30]. Each node in the augmented OER ontology indicates a micro OER chunk. A chunk is the smallest unit in the micro learning settings, normally a fine-cut piece of an OER from its provider, and it has an apparently shorter time length (preferably less than 15 mins) than its original shape. It can be a mini concept or knowledge point, tinier than what the teachers used to deliver; or it can be a cut of course video or lecture notes, or a course settings come along with a concept, such as assessment, task, reading material and so on [20].

There is not any totally independent chunk and each of them is part of a relational web rather than merely a conceptual object [31]. This ontology is used to explicitly classify the OERs to recommend among a pedagogically defined set of distinctive main concepts, fed as the raw material in the reasoning process of MLaaS [27] [31].

A conceptual graph of the augmented OER ontology is shown as Figure 2.

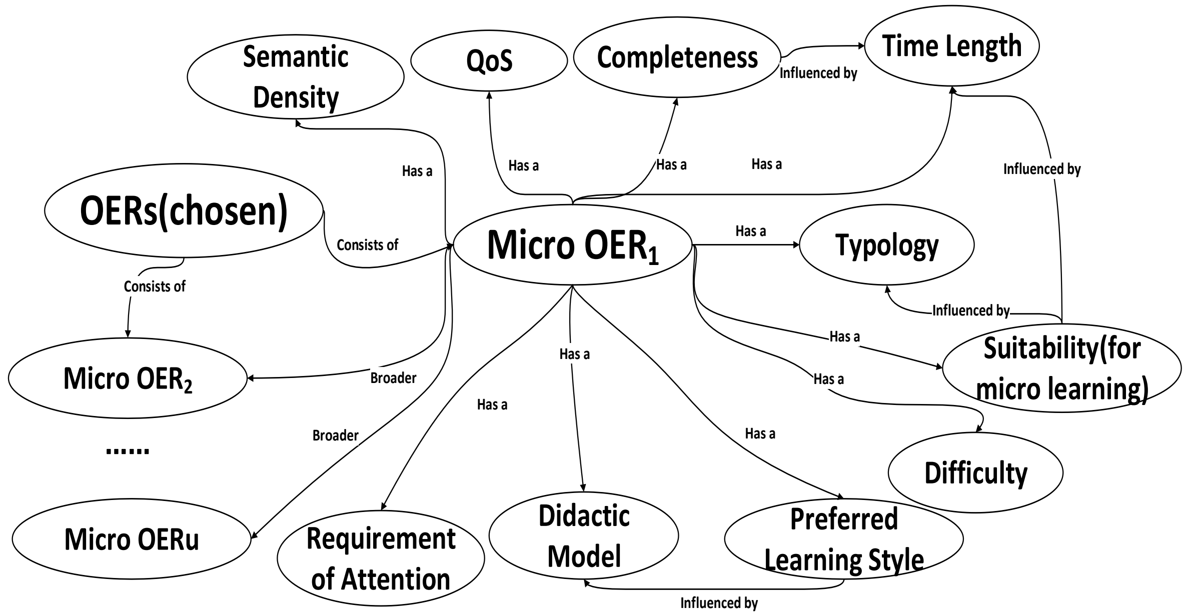

Figure 2. Conceptual Graph of Micro OER ontology

4.1.2 Augmented Micro Learning Learner Profile Taxonomy

From the user-based view, a main ontology, on which all learner profiles are based, is named as the *Benchmark* ontology, where the element *Learner* is put at the center of the graph [27]. Acting as the instances of the presetting domain ontologies, a specific learner profile oriented to micro learning is a set of nodes from the *Benchmark* ontology versus a node in the augmented micro OER ontology. It contains plenty of annotations in terms of their learning behaviors and context. A conceptual graph of the benchmark ontology is shown as Figure 3.

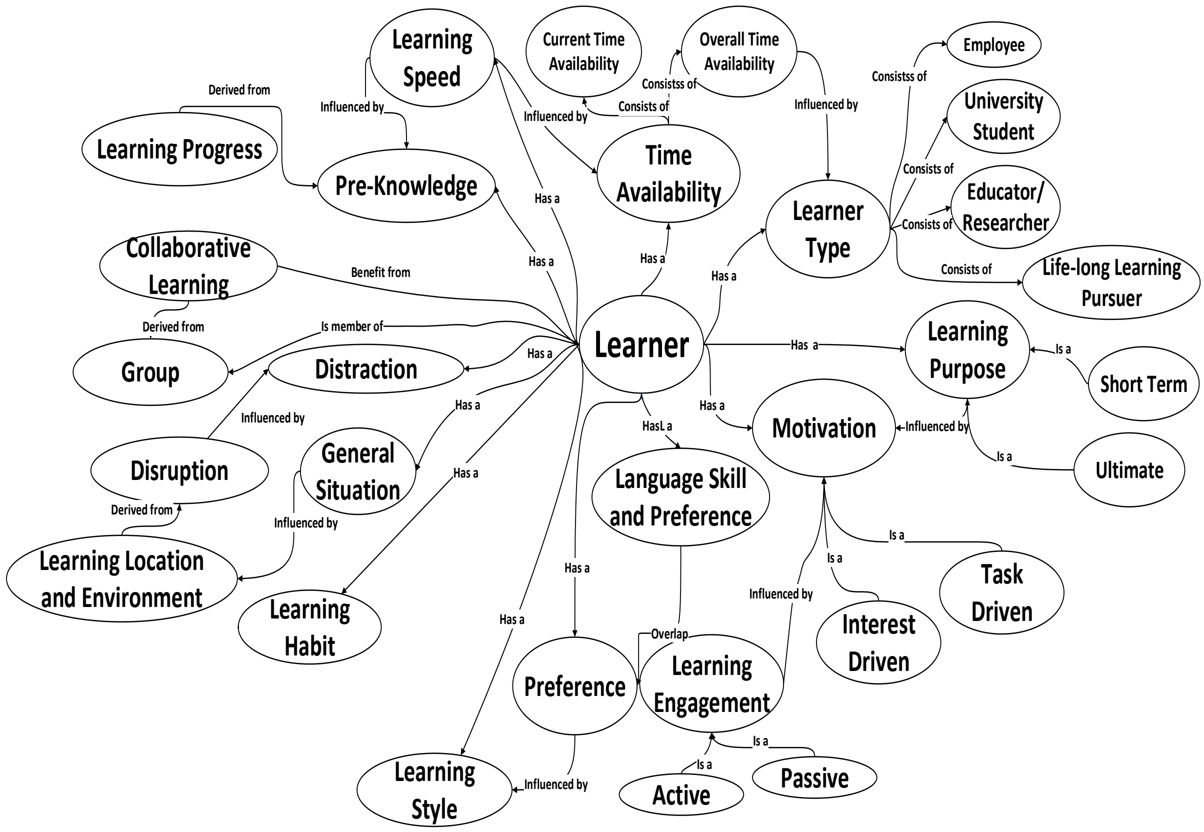

Figure 3. Concept Graph of Benchmark Ontology for a Learner Profile in Micro Open Learning

4.2 EDM and LA Strategy

Reporting learners' data visually and statistically to demonstrate their unique learning story and also their learning constraints (such as time availability) is crucial. This plays a significant role in assessing learners' study status, estimating learners' study progress, and carrying out learning strategic decision-making. It is responsible for the benchmark setting for the routine data extraction from open learning platform.

For the bottom level (i.e. data level) of the knowledge base, to technically operate the semantic learner profile and knowledge base construction for micro open learning, data filled in the graphs come from two sides, the explicit data collection (e.g. through mandatory requests) and implicit data tracking (e.g. automatic extraction) [31].

In addition, rather than developing the domain ontology for OERs by ourselves, a general structure of courseware ontology is built jointly by making use of existing ontologies, which had been extracted from main OER providers, such as universities involved in major open courseware alliances (e.g. participating institutions in edX[1]), or from the Linked Open Data Cloud community[2] [32].

The investigation over 'big' open learning data comes up to the OER side. Among the massive OERs, three types of relations are mainly targeted to be foreseen:

- ConsisitsOf is an inclusion relation. This relation can be generally found between two OERs or one OER and one micro OER. Two items with this relation are located in different hierarchies of the augmented micro OER ontology.
- RequiredSequence is a strong order between two items (OER or micro OER), where the former micro OER is necessarily to be learnt before the latter one, due to course setting and educational consideration.

---

[1] https://www.edx.org/schools-partners

[2] http://lod-cloud.net/

- RecommendedSequence is a weak order relation between two items (OER and micro OER), where the former micro OER is suggestively to be learnt before the latter one, according to the instructors' guidance, but it is not mandatory.
- Certainly, two items (OER or micro OER) can have no relation at all.
- Both relations regarding sequence can be inherited by entities' descendants, for example, if there is a RecommendedSequence(R_1, R_2) indicating an OER R_1 is preferably learnt prior to R_2, then, for $MR_1 \in MR_2$, there is a RecommendedSequence(MR_1, MR_2)

The purpose of the EDM/LA is to amend, enrich and validate the aforementioned ontologies built manually and extracted semi-automatically, and verify and weigh the importance of discovered relations. Our EDM/LA is combinedly realized by two parts [33], based on on-campus mobile learning data (i.e. structured data) and 'big' open learning data (i.e. unstructured data). In particular, we are carrying out the experimental EDM/LA by substantially analyzing the real data of learning behaviors of students from a public university in Australia. The data are collected from the main learning management system (LMS) and data warehouse of the university. This analysis aims to identify the regular patterns of students getting involved in blended learning (i.e. on-campus learning and e/m learning), for example, whether and how often they adopt micro learning mode to accomplish learning tasks and to explore the major factors that affect their learning habits and, most importantly; and the rules of how features listed in the personalized learner model mutually affect and interrelate, and act on their learning outcomes. In this stage, we are discovering the potential trends, which cannot be directly shown from the data we can achieve, and then we apply such findings into open learning scene and infer what are behind the scene. The detailed data sets are illustrated as follows:

Table 1. EDM/LA Data Sources from University Warehouse

Data type	Purpose
Learners' exact time of logon/out for each time	To know how long they stay online for each time
The IP address or gateway information of their internet connection	To know their exact learning location and surroundings
Mobile device information, mobile operator information and mobile OSs	To know their general situation
Their personal enrollment information (full time or part timer nationality)	To know their learning time availability, organization and language skills
Their residential information (session address and permanent address)	To understand their distances to campus and the potential modes of transportation they adopt)
Subjects they have chosen (current)	To know their academic background and field
Subjects they have chosen (historical)	To know their academic background and field
Historical grades	To know their academic background and infer pre-knowledge level
Course materials they have accessed (material type, topic, length, requirement associated with them)	To know their learning habits (how they prefer the learning resources to be passed on)
Course requirement/milestones set in LMS (by instructor)	To know the suggested learning schedule
Their detailed learning activities (What they do when staying online and how long they spend on each specific learning activity, type of resource they access for each specific time)	To know their learning habits, learning engagements, learning speed and so on.
Their interactions with LMS and learner-generated content (from forum and thread, etc.)	To know their preference, interests and measure their engagement.

Frequencies of they participate in interactive learning activities (e.g. forum, thread)	To know their engagement
Extent of completeness for each learning activity	To know whether they finished an entire step of learning or drop off halfway
The learning paths they have gone through (the sequence of they access learning resources over LMS)	To further establish optimal learning paths
Their learning achievement (grades and final marks if possible)	To know how their learning behaviors affect their learning outcomes
Groups or teams they have participated in	To know their collaborative learning performance and similarities/changes of learning time frame among learners

The study is subsequently extended and applied into a larger scale, by analyzing 'big' data from real open learning activities. Data mining means with different aims are concluded in the first column in the Table 2.

Table 2. EDM/LA Scheme for Open Learning Data

Technique	**Object**	**Purpose**
Prediction	well-defined micro OERs	to establish a recommendation model for students in similar situations in the future
Structure Discovery	well-defined micro OERs	for web documents using clustering methods in order to personalize e-learning based on maximal frequent item sets
Latent Knowledge Estimation	Non-micro OERs	to discover which stages of them are generally finished within relatively larger time length
Structure Discovery	Non-micro OERs	to determine time spans where the pauses made by learners usually fall in
Factor Analysis	Non-micro OERs	to find out the actual reasons why learners spent more time on these stages and made such pauses
Latent Knowledge Estimation	Non-micro OERs	to measure potential suitability of micro learning (from learners' frequencies of using fragmented time pieces)
Factor Analysis	Non-micro OERs	to identify resources' suitability for micro learning, for example, whether a hands-on practice is needed, or whether the OER delivery is necessarily associated with lots of writing or computation work which is inconvenient to complete on mobile devices
Prediction	Subscription OERs	to determine when to push information to learners in the best timing and remind them
Clustering	All micro OERs	to determine their correlations for better repository purpose
Relationship Mining	Time Availability	to discover the correlation between their overall time availability and learners' types
Clustering/Prediction	Time Availability	to involve similar learners into cohorts and build potential time frame for them overall learning schedule
Latent Knowledge Estimation	Learning habit (learning time distribution)	to discover whether there are regular patterns of time organization within time frame among learners in or across cohorts
Latent Knowledge Estimation	Learners' latest learning contents and activities	to retrieve back and profile learners' learning recentness
Categorization	Learning habits	to set up a unique learning habit summary for each learner
Relationship Mining	Learners' learning location data	to know the degree of distraction and how it interrelates to disruption from external environment,

Relationship Mining	Learners' mobile app usage	to know the degree of distraction and how it interrelates to disruption from the content on mobile internet
Social Network Analysis	OERs in affiliated social networks	to distinguish information that can be useless, harmful and may cause time wasted for learners.
Social Network Analysis	Other content in affiliated social networks	to screen well-recognized information in order to recommend to learners as their learning augmentation besides the OERs (text mining technique employed)

To a wider extent, the establishment of data level can involve integrating heterogeneous OCW repositories, refining and blending available OERs into micro learning context and publishing their metadata as linked data. Because in recent years some educators and researchers have made great efforts on publish and popularize the OER in terms of the linked data concept, a workflow developed with this extended aim is generally divided into six phases:

1. Identify and select heterogeneous data sources to determine the scope of the content.
2. Model vocabularies for OER domain.
3. Data extraction.
4. Generate standardized data description (e.g. RDF data).
5. Publish linked data.
6. Consume and display linked data.

5. Ontological Approach for Cold-start Problem

5.1 Representation of Learner Profile

Adopting ontologies as the basis of the learner profile is crucial in addressing the cold start problem in micro OER delivery. It allows the initial learner behavior to be matched with existing and pre-known knowledge in the ontologies and relationships among them.

The learner profile is managed by MLaaS by two parts: static part and dynamic part. The static part can be represented by a vector, which contains the demographic and educational information. By matching these two augmented ontologies, respectively for item and user, the dynamic part of a learner node is denoted as a pair, $L_j = \{MR_u, ML_j\}$, $L_j \in L$. Herein, the element MR_u denotes the u^{th} micro OER, as introduced in the Section 4.1.2, which is a particular version of the micro OER ontology, and a three dimensional element $ML_j\{P_{u,j}, TA_j, D_j\}$is exclusive to j^{th} learner during the micro learning process. Herein, the element $P_{u,j}$ indicates the learner's preference, TA_j indicates the j^{th} learner's instantly time availability, and D_j denotes the level of distraction in terms of the given learning environment and surroundings.

Whenever MLaaS gathers any information from the learner's learning process over OER, the learner profile will be updated in regards to ML_j.

5.2 Preference Propagation

Provided the cold start condition for the first micro OER delivery, a learner is required to quickly mark down a preference on a specific micro OER. Consequently, a spreading activation approach is applied to maintain the preference against its parent node (i.e., the R_v is the v^{th} OER where the MR_u derived from) as well as updating learner profile. It propagates the learners' preference upwards the hierarchy of micro OER ontology based on activation values. In other words, the preference has been obtained from a micro OER to its ancestor and spread in its superclass (i.e., OER) level. An example of the spreading activation is shown in Figure 4.

A partial view of augmented micro OER ontology in 'information technologies' area is shown in Figure 4. Particularly, it describes an 'e-business' OER from an Australia provider, OpenLearning[3]. At the bottom level of the ontology, nodes which are depicted with oval shape

[3] https://www.openlearning.com/

typically conform to the standard of micro OER. Red integers shown in nodes with rectangle shape are preference values from a learner versus target OERs. The algorithm 1 is proposed to execute the process of preference propagation.

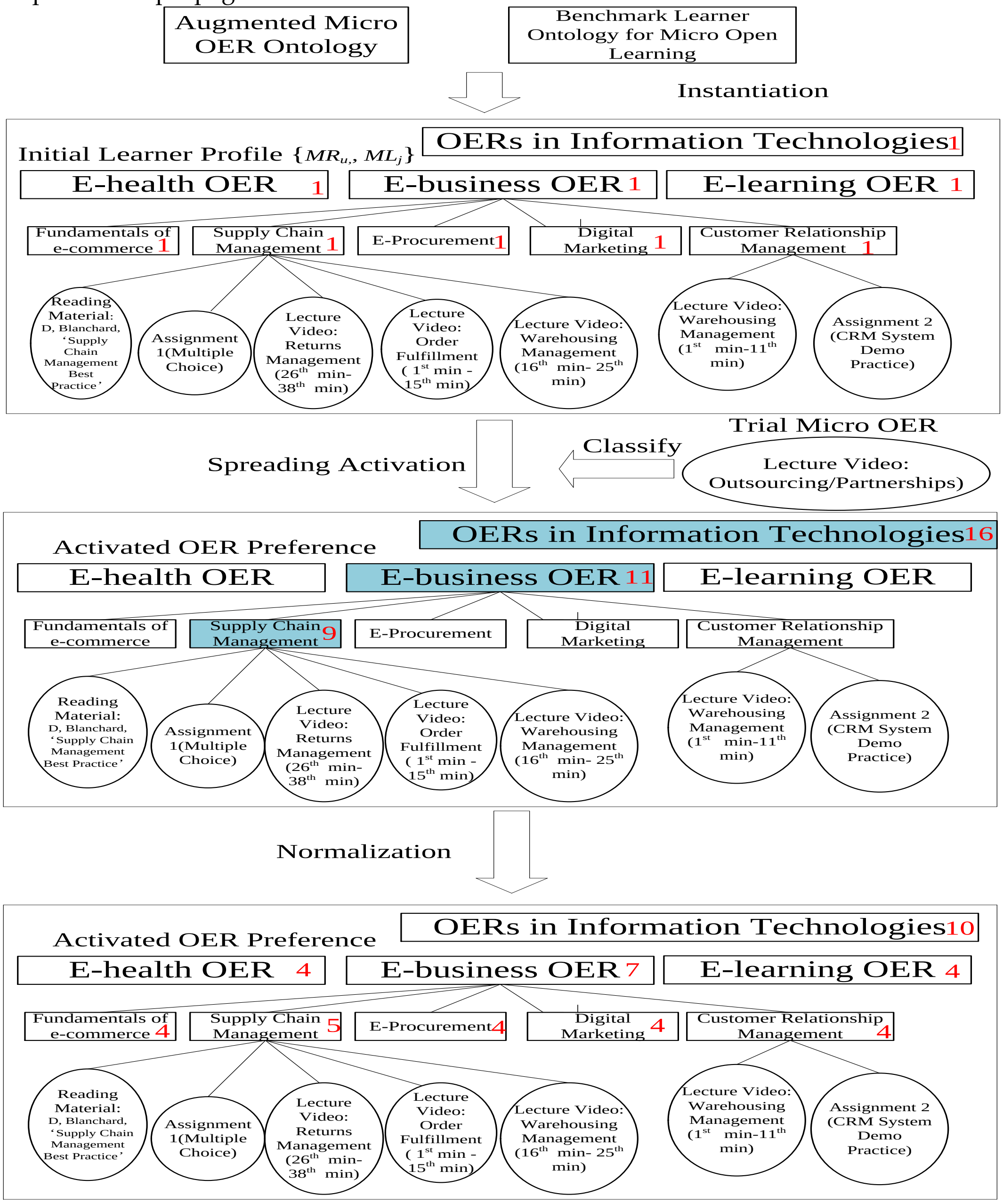

Figure 4. Partial View of the Augmented Micro OER Ontology and Spreading Activation for a Learner's Preference on OER

Algorithm1: Preference Propagation

Input: *Dynamic part of learner profile L_j ={ MR_u, ML_j }, a trial micro OER MR_u , $L_j \in L$*

Output: *Updated dynamic part of learner profile with updated $P_{u,j}$ value in the triple dimensional set ML , $P(R_v)$and $Activation(R_v)$, preference value and activation value for the OER R_v*

//Step 1: Spreading Activation

begin: *Initialize PriorityQueue;//PriorityQueue is the set of OERs within the same discipline where R_v belongs to*

Set Activation of all micro OER to 0

for each $L_j \in L$ ***do***

if *($MR_u \in R_v$)* ***then***

$Activation(R_v)= P(R_v)$

PriorityQueue.Add(R_v)

```
                end if
    end for
    while PriorityQueue.Count >0 do
        Sort PriorityQueue; //activation values in descending order
        Select the first item MRu  in PriorityQueue // Rv with highest P value
            Remove Rv in PriorityQueue
        for each Rv do
LinkedOERs=GetLinkedOERs(Rv)//get linked nodes of Rv
            for each Rw in LinkedOERs do //propagate activation to its neighbors
                Activation(Rw)+=Activation(Rv)*Weight(Rv, Rw )
                PriorityQueue.Add(Rv)
                Sort PriorityQueue
            end for
        end for
    end while
            //Step2: Learner Profile Normalization
    for each   Lj ∈L do
P(Rv)= P(Rv)+ Activation(Rv)
k = 1/ sqrt( sum_{t=1}^{v} (Rv)^2 )   //normalization factor
P(Rv)= P(Rv)*k //normalization
Pv,j=P(Rv)
    end for
end
```

The normalization factor acts on preventing the propagated preferences from escalating continuously to such an extent that exceeds a reasonable range, which could result in difficulty of data processing in the forthcoming process. The confidence degree for the propagated preference of OER is recorded as $CD(P_{v,j})$.

5.3 Instant Time Availability

The system is able to obtain explicit information on how long the learner can (or would like) to spend on a micro OER through mobile devices in the real time. As a mandatory request, a learner is required to input his instant time availability at the beginning of every micro learning activity.

According to the system setting, suggestively the instant time availability, TA_j, is represented by an integer from 1 to 15. However, if the learner is not pretty sure how long he is able to spend on the micro OER at once, he is free to leave a time span, which can be continuous integers in the same range.

5.4 Learner Feature Prediction

5.4.1 Demographic Classification

In our prior work [20], we have discussed the key issues that might cause distraction in micro open learning, which generally came from two sides, the social side and environmental side.

In addition, MLaaS investigates existing learners' degree of distraction as reference, and senses every learner's location information through built-in functions in mobile devices. Based on the given taxonomy and augmented ontology, we carry out a demographic classification that aims to cluster learners into cohorts, in order to match them with micro-pieces of OERs.

The mechanism of classification is designed as, learners who have similar static information, involving employment and/or education background, occupation, and similar learning environment/location, are more likely to face similar level of distraction. For the same reason, their overall time availabilities would more likely fall in the same range. Herein MLaaS tries to associate a learner into a pre-clustered learner group, by applying the stereotyping technique to fulfill the requirement of demographic classification.

For a newly joined learner, L_j, an ensemble method of a binary classifier and a one-against-all model is utilized to obtain multi-class classification achievements [34][35], in order to predict its category, C_j. The system is trained with existing set of learners, L. Typical binary classification techniques, i.e., C4.5 decision tree [36][37] or Naive Bayes classifier [38] can be employed to serve as the base algorithm (i.e. training algorithm F in Algorithm 2) in order to produce a suitable classifier, CF_k. A new learner L_j is classified with the label k, whose CF_k produces the highest value of $\hat{y}$.

Algorithm2: New Learner Classification （one-against-all）

Input: *Sample (Current learner set L), Lables y (where y_i is the label for a sample learner L_i and $y_i \in \{1,2,\ldots K\}$), Training Algorithm F, a new learner L_j*

Output:*Category of the new learner L_j, C_j*

begin:
    for each *k in {1, 2, …K }* ***do***
      set a new label vector z_i for y_i,
        if *(y_i=k)* ***then***
            z_i=1
          else
            z_i=0
        end if
          $C_k=F(L,z_i)$ //use binary classification technique to produce classifiers
    end for
    for *$L_j \in L$* ***do***
        $\hat{y} = \arg \max_{k \in \{1,2 \ldots k\}} CF_k(L_j)$
    end for
    output C_j =k //category of the new learner L_j
end

Given L_j is categorized into the C_k, afterwards, the learner's neighborhood, NB_j, is calculated by the Algorithm 3. This aims to match a new learner's category with an existing learner's category.

Algorithm3: Neighbourhood Calculation

Input: *new learner set N and Existing learner set L*

Output: *set of neighbours, NB_j, of a new learnerL_j*

begin: *Build a binary classifier*
    Execute the one-against-all model//as in Algorithm 2
    Build the ensemble method of multiclass classifier//categorize new learners
    for each *L_j in N* ***do***
      NB_j=null
      Predict C_j// L_j's category
        for each *L_i in l* ***do***
          Retrieve C_i
          if *C_j== C_i* ***then***
            NB_j.add(L_i)
          end if
        end for
    end for
end

Hence, the demographic classification is realized according to learners' static and location information. Once new learners join into the open learning scenario, MLaaS responds immediately to classify them into clusters.

5.4.2 Similarity Measure between Two Learners

MLaaS is responsible to find the similar existing learners in the discovered demographic categories, so as to recommend them micro OERs that were recognized as suitable to learn in a given time span, situation and environment.

Learners' learning location information is sensed from the location service embedded in the mobile devices. Thus, the similarity between two learners, L_i and L_j, is evaluated using the equation (1).

$$sim(i,j) = [(\sum_{l=1}^{m} S_l W_l)^2 + (SLo_{i,j} W_{i,j})^2]^{1/2} \quad (1)$$

where S_l is the similarity value of the l^{th} attribute in the static part of learner profile and the W_l is its corresponding weight. $SLo_{i,j}$ denotes their similarity on location and $W_{i.j}$ denotes the weight for location factor.

5.4.3 Distraction Prediction

Thus, in terms of the equation (2), the distraction value can be estimated in accordance with the action that any member in a same cluster indicates the predicted distraction level.

$$D_{j,Lo_a} = \frac{1}{2} * \left(\frac{\sum sim(i.j) \bullet d_{i,Lo_a}}{\sum sim(i,j)} + d_{j,Lo_a}\right) \quad (2)$$

where $d_{j,Loa}$ is the self-identified degree of distraction the learner L_j felt in the location Lo_a, acquired by mandatory request. This follows the expectation that the learners who have similar general situation (i.e. social factors) and surroundings (i.e. environmental factors) are in high probability to have similar degree of distraction.

The confidence degree for the predicted distraction is depicted as $CD(D_{i,Loa})$.

5.5 Integration of Recommendation Results

5.5.1 Downwards Propagation

In the Section 5.1 we have merely obtained the preference of a learner on an 'entire' OER rather than on a micro OER, now the preference values are again propagated downwards the ontology hierarchy. Consequently, each micro OER node receives an estimated preference value from its ancestor. This propagation process is executed with a decay factor. For each micro OER the final preference value, $P_{u,j}$, can be calculated use the following equation (3).

$$P_{u,j} = \frac{\sum P_{R,j} CD(P_{R,j}) + \sum P_{v,j} CD(P_{v,j}) Q(u,v)}{\sum CD(P_{R,j}) + \sum CD(P_{v,j}) Q(u,v)} \quad (3)$$

where R is the set of all the nodes in the higher hierarchy than MR_u, R_v is a direct ancestor of node MR_u and $Q(u,v)$ depends on the count of level between MR_u and R_v.

$$Q(u,v) = L(|uv|) = \begin{cases} L(0) = 1 \\ L(l) = (1-\lambda) L(l-1) \end{cases} \quad (4)$$

and the confidence degree for the descendant node, in regards to the $P_{u,j}$, is calculated as the average of the confidence values in its ancestors, decreased by a decay factor, μ.

$$CD(P_{u,j}) = \frac{\sum CD(R)}{|R|} - \mu \quad (5)$$

As far as all values of the three attributes, denoting preferences, instant time availability and degree of distraction, in the set *ML* are settled, a complete learner profile is constructed from the initial little-known information by the MLaaS.

5.5.2 Micro OER Sorting Rules

For each micro OER, once MLaaS has acquired its final preference value and confidence degree, those nodes which do not meet the minimum requirement of confidence degree is rejected by the system.

To generate a list of recommended micro OERs, where the ones with higher learners' interests are placed at the top. For two micro OERs MR_u and MR_w, their sequence is determined according to some heuristic rules which are defined in accordance with the extraction of three kinds of relations discussed in the Section 4.1.1. These rules are executed sequentially with priority.

1. If there is a RequiredSequence relation between these two micro OERs, the prerequisite one is placed above (refer to the Section 4.1.1).
2. If the preference regarding these two OERs, $P_{u,j}$, $P_{w,j}$, the former one is higher than the latter one, then the MR_u is above MR_w
3. If, in the absolute terms, the confidence degree $CD(P_{u,j})$ is high and the $CD(P_{w,j})$ is low, then the MR_u is above MR_w.
4. If there is a RecommendedSequence relation between these two micro OERs, the one which is suggested to be accessed first is placed above (refer to the Section 4.1.1).
5. The micro OER which is more related to the learners' education background, or falls in the relevant disciplines or inter-disciplines is placed with priority if the disciplinary difference between these two candidate micro OERs is obvious.
6. Otherwise the recommended micro OER list is randomly ordered if none of the above rules applies.

Herein, the first rule is deemed as a hard rule which should be strictly obeyed and the rest rules are soft rules which can be violated with educational consideration, from case to case.

5.5.3 Recommendation Results Optimization

MLaaS consumes the value *P* and *D* in conjunction with their *TA* to compare with the attributions and requirements annotated in the metadata of the augmented OER ontology.

The next step is to integrate the outcomes from the Section *5.4*, a fitness function will convert these selected multidimensional arrays into one variable. Hence, this problem is hereby properly transferred to a multi-objective optimization problem.

To initiate the constrained multi-objective optimization, candidate learning path solutions (chromosomes) are randomly generated where each of them is a learning path with a series of micro OERs. For a chromosome, its violation degree is investigated by examining the relations between each contiguously prior/posterior micro OER pair against the first 5 rules listed in the Section 5.4.2, and then summing up. For a such pair in a chromosome, its violation degree, $VD(MR^t, MR^{t+1})$, is calculated by the weighted sum of its violations against rule 2 to rule 5, respectively, where MR^t is the t^{th} micro OER in *k* and MR^{t+1} is the $t+1^{th}$. The higher the violation degree is, the more serious the candidate learning path violate the rules. The violation degree of a candidate learning path, *k*, is calculated using the following equation (6)

$$VD_k = \sum VD(MR^t, MR^{t+1}) \qquad (6)$$

Thereafter, let the variable RA_u denote the degree of required attention of a given micro learning resource, MR_u, whose real-time suitability for micro learning, $RT_{u,j}$, is calculated by comparing with the learner, L_j's predicted distraction, using the following equation (7):

$$RT_{u,j} = \{(RA_u)^2 + [CD(D_{j,Lo_a}) * D_{j,Lo_a}]^2\}^{1/2} \qquad (7)$$

Hence, for the candidate learning path, k, $RT_{k,j}$ denotes the sum of the real-time suitability of micro OERs it contains. Similarly, $P_{k,j}$ sums up all the predicted preferences from the learner L_j versus micro OERs that k contains.

$$\eta = \min(\alpha VD_k + \beta RT_{k,j} + \gamma / P^1_{k,j} + \delta / P_{k,j}) \tag{8}$$

where α, β, γ and δ serve as weight for each variable and suggestively $\alpha>\beta>\gamma>\delta$, $P^1_{k,j}$ denotes the L_k's preference value of the first micro OER in the candidate learning path k.

The algorithm 4 indicates typical steps of make the first recommendation

Algorithm4: Micro OER Recommendation in a Cold Start Condition

Input: $P_{u,j}$ *(the Learner L_k's predicted reference to the micro OER MR_u), $D_{j,Loa}$ (predicted distraction level), $CD(P_{u,j})$ and $CD(D_{i,Loa})$ (their confidence degree), RA_u (the degree of required attention of MR_u), TA_j (the instant time availability), rules (1st-6th)*

Output: *the tag of a micro OER which acts as the first delivery*

begin: *Randomly generate candidate learning paths as chromsomes*
    for each *chromosome k* ***do***
      Select micro OER it contains
        for each *MR_u in a chromosome k,*
            Caculate its $P_{u,j}$ and $CD(P_{u,j})$.
          Import $D_{j,Loa}$, $CD(D_{i,Loa})$ and RA_u
          Caculate its $RT_{u,j}$
        end for
      Calculate k's VD_k
      Use equation (8) to evaluate its fitness η
    end for
    while *iteration times < max iteration time* ***do***
    apply heuristic approach to generate new candidate solutions
      for each *new chromosome k'* ***do***
          check time length of the first micro OER in k', $TL^1_{k'}$
        if *$TL^1_{k'}$ is in the range of TA_j*
        keep k'
        otherwise
        reject k'
        end if
          evaluate the fitness of k', η, using equation (8)
      end for
    replace chromosomes with higher η values
    end while
    output the selected chromosome k''with minimum η and satisfied $TL^1_{k''}$
    select the first micro OER in k''as the first delivery
end

By this means, the heuristic algorithm 4 infers a suitable micro OER as the first attempt of learning resource recommendation in the novel open learning experience through MLaaS.

Along with the successful launch of solution to the well-known cold start problem in micro learning, learners' upcoming behaviors will be continuously acquired by MLaaS to feed the reasoning engine.

Conclusions

In this paper, we introduced a study aimed at dealing with the adaptive micro OER delivery. We proposed a tailored system for this MLaaS. Besides its technical details and working principles, we primarily focused on the knowledge base construction. It was built using a top-down approach, by having ontologies at the pattern level first. Using this, a strategy on processing data was then developed at the lower level. This supported the decision-making process of the micro OER recommendation system. However, because both the system and user are new, learner's information deficit in MLaaS delayed the commencement of adaptive micro open learning and MLaaS operation. A detailed approach was therefore provided to deal with this so called 'cold start problem' based on predicting learners' features from the initial little-known information.

Our future work will extend the EDM/LA work, and prototyping this ontological approach for cold start problem and further developing its corresponding component in MLaaS. This will be further evaluated by measuring the prediction accuracy. We will also engage real learners to compare the quality of recommendation. Apart from the 'new user' cold start problem discussed in this paper, we will look for solutions to deal with 'new items', i.e., micro OERs, by using a queue-jumping method to insert them into established learning paths.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. J. Hilton, 'Open Educational Resources and College Textbook Choices: a Review of Research on Efficacy and Perceptions', Educational Technology Research and Development, vol.64, no. 4, pp.573-590, 2016.
2. J. Baggaley, 'MOOC Rampant', Distance Education, vol. 34, no. 3 pp. 368-378, 2013.
3. J. Hylen, D.V. Damme, F. Mulder and S. D' Antoni, 'Open Educational Resources: Analysis of Responses to the OECD Country Questionnare', OECD Education Working Papers No. 76, June 2012.
4. I. Nawrot and A. Doucet, 'Building Engagement for MOOC Students: Introducing Support for Time Management on Online Learning Platforms', Proceeding of WWW'14 Companion, 2014
5. I. Warrd, A. Koutropoulos, N. O. Keskin, S.C. Abajian, R. Hogue, C. O. Rodriguez and M. S. Gallagher, 'Exploring the MOOC Format as a Pedagogical Approach for m-learning', the 10th World Conference on Mobile and Contextual Learning, Beijing, China, October 2011.
6. G. Sun, T. Cui, S. Chen, W. Guo and J. Shen, 'MLaaS: A Cloud Based System for Mobile Micro Learning in MOOC', 4th IEEE International Conference on Mobile Service (MS), New York, US, July 2015, pp.120-127.
7. G. Sun, T. Cui, K. Li, D. Xu, S. Chen, J. Shen and W. Guo, 'Towards Bringing Adaptive Micro Learning into MOOC Courses', the 15th IEEE International Conference on Advanced Learning Technologies (ICALT), Hualien, Taiwan, July 2015.
8. V. P. Dennen and S. Hao, 'Intentionally Mobile Pedagogy: the M-COPE Framework for Mobile Learning in Higher Education', Technology, Pedagogy and Education, vol. 23, no. 3, pp.397-419.
9. A. Trifonova and M. Ronchetti, 'Mobile Learning: Is Anytime + Anywhere = Always Online?' 6th IEEE International Conference on Advanced Learning Technologies, Kerkrade, Netherlands, July, 2006.
10. J. Attewell, a Technology Update and m-Learning Project Summary, Mobile Technology and Learning, Learning and Skills Development Agency, UK, 2005.
11. H. Wu, L. Hamdi and N. Mahe, 'Tango: A Flexible Mobility-enabled Architecture for Online and Offline Mobile Enterprise Applications', 11th International Conference on Mobile Data Management, Kansas City, MO, USA, May, 2010.
12. M. J. W. Lee and C. McLoughlin, Teaching and Learning in the Web 2.0 Era: Empowering Students through Learner-Generated Content, International Journal of Instructional Technology and Distance Learning, vol. 4, no. 10.
13. D. Kovachev, Y. Cao, R. Klamma and M. Jarke, 'Learn-as-you-go, New Ways of Cloud Based Micro-learning for the Mobile Web', 10th International Conference on Web-based Learning, Hongkong, December 2011.
14. P.A.Bruck, L.Motiwalla and F. Foerster, 'Mobile Learning with Micro-content: A Framework and Evaluation', 25th Bled eConference, Bled, Slovenia, 2012, pp.527-542.
15. T. Hug and M. Lindner, 'ML: Emerging Concepts, Practices and Technologies after e-Learning', Proceedings of micro learning 2005, Austria, June 2005Z. Sun and J. Shen, A high performance peer to cloud and peer model augmented with hierarchical secure communications, Journal of Systems and Software, vol. 78, no. 7, pp.1790-1796.
16. M. I. Souza and S. F. D. Amaral, 'Educational Micro content for Mobile Learning Virtual Environments', Creative Education, vol. 5, pp. 672-681, 2014.

17. Z. Sun and J. Shen, A high performance peer to cloud and peer model augmented with hierarchical secure communications, Journal of Systems and Software, vol. 78, no. 7, pp.1790-1796.
18. G. Sun, T. Cui, W. Guo, G. Beydoun, D. Xu and J. Shen, 'Micro Learning Adaptation in MOOC: a Software as a Service and a Personalized Learner Model', the 14th International Conference on Web-based Learning (ICWL), Guangzhou, China, November 2015, pp.174-184.
19. L. Wang and J. Shen, 'Multi-phase ant colony system for multi-party data-intensive service provision', IEEE Transactions on Service Computing, vol. 9, no. 2, pp. 264-276, 2016.
20. G. Sun, T. Cui, J. Yong, J. Shen and S. Chen , 'MLaaS: a Cloud-based System for Delivering Adaptive Micro Learning in Mobile MOOC Learning'. IEEE Transactions on Services Computing, online first http://dx.doi.org/10.1109/TSC.2015.2473854
21. R.R.Starr and J. M. P. de Oliveira, ,'Concept Maps as the First Step in an Ontology Construction Method', Information Systems, vol. 38, no. 5, pp. 771-783, 2013.
22. J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh and A.H. Byers, Big data: The next frontier for innovation, competition, and productivity, 2011.
23. C. Romero and S. Ventura, 'Educational Data Mining: a Review of the State-of-the-Art', IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol.40, no. 6, pp. 601-618, 2010.
24. A. Sieg, B. Mobasher and R. Burke, 'Ontology-Based Collaboraitve Recommendation', the 8th Worksho on Intelligent Techniques for Web Personalization and Recommender Systems, Hawaii, USA, June 2010, pp. 20-32.
25. S. Miranda, G. R. Mangione, F. Orciuoli, M. Gaeta and V. Loia, 'Automatic Generation of Assessment Objects and Remedial Works for MOOCs', 12th International Conference on Information Technology Based Higher Education and Training, Antalya, Turkey, Octorber 2013.
26. B. Lika, K. Kolomvatsos, S. Hadjiefthymiades, 'Facing the Cold Start Problem in Recommender Systems', Expert Systems with Applications, vol. 41, no. 4, pp. 2065-2073, 2014.
27. G. Sun, T. Cui, G. Beydoun, S. Chen and J. Shen, 'Profiling and Supporting Adaptive Micro Learning on Open Education Resources' the 4th International Conference on Advanced Cloud and Big Data (CBD), Chengdu, China, August 2016, pp.158-163.
28. E. Cambria, D. Olsher and D. Rajagopal, 'SenticNet3: a Common and Common-sense Knowledge Base for Cognition-driven Sentiment Analysis', the 28th AAAI Conference on Artificial Intelligence (AAAI), Quebec, Canada, July 2014, pp.1515-1521.
29. D. Dou, H. Wang and H. Liu, 'Semantic Data Mining: a Survey of Ontology-based Approaches', the 9th IEEE International Conference on Semantic Computing (ICSC), California, USA, February 2015, pp. 244-251.
30. S. Miranda and G. Albano, 'Personalized Learning in Mathematics', Journal of E-learning and Knowledge Society, vol. 11, no.1, pp.25-42, 2015.
31. A. Moreno, A. Valls, D. Isern, L. Marin and J. Borras, 'SigTur/E-Destination: Ontology-Based personalized Recommendation of Tourism and Leisure Activities', Engineering Applications of Artificial Intelligence, vol.26, no.1, pp.633-651, 2013.
32. N. Capuano, L. Dell'Angelo, F. Orciuoli, 'Ontology Extraction from Existing Education al Content to Improve Personalized e-Learning Experiences', 3rd IEEE International Conference on Semantic Computing, Berkeley, USA, September 2009, pp. 577-582.
33. P. Markellou, I. Mousourouli, S. Spiros and A. Tsakalidis, 'Using Semantic Web Mining Technologies for Personalized e-learning Experiences', Proceeding of the Web-based Education, Grindelwald, Switzerland, pp. 461-826, 2008.
34. J. Milgram, M. Cheriet and M. Sabourin, 'One against One or One against all: Which One is Better for Handwriting Recognition with SVMs', 10th International Workshop on Frontiers in Handwriting Recognition, France, October 2006.
35. M. Galar, A. Fernandez, E. Barrenechea, H. Bustince and F. Herrera, ' An Overview of Ensemble Methods for Binary Classifiers in Multi-class Problems: Emperimental Study on One-vs-One and One-vs-All Schemes, Pattern Recognition, vol. 44, pp.1761-1766, 2011.
36. S.B.Kotsiantis, 'Supervised Machine Learning: a Review of Classification Techiniques', Informatica, vol.31, no. 3, pp. 249-268, 2007.

37. K. Polat and S.Gunes, ‘a Novel Hybrid Intelligent Method Based on C4.5 Decision Tree Classifier and One-against-All Approach for Multi-class Classification Problems’, Expert System with Applications, vol.36, pp.1587-1592, 2009.
38. L. A. Dalton and E. R. Dougherty, ‘Optimal Classifiers with Minimum Expected Error within a Bayesian Framework-Part II: Properties and Performance Analysis’, Pattern Recognition, vol. 46, no. 5, pp. 1288-1300, 2013.
39. J. Shen, G. Beydoun, S. Yuan and G. Low, ‘Comparison of bio-inspired algorithms for peer selection in services composition’. In: Proceedings of IEEE international conference on services computing, Washington, DC, pp 250–257, 2011.
40. D. Xu, C. Wijesooriya, Y.-G. Wang, G. Beydoun, ‘Outbound logistics exception monitoring: A multi-perspective ontologies approach with intelligent agents’, Expert Systems with Applications, 38 (11), pp. 13604–13611, 2011.